HCC Grid Applications

Derek Weitzel & Adam Caprez

HCC Grid Introduction

- Use GlideinWMS for grid submission
  - CMS Factory at UCSD

- Test applications on local clusters before going to grid

- Most workflows are designed by HPC/HTC application specialists.

HCC Grid Introduction

- Run at 24 sites:
    - All CMS T2's and CMS T1
    - Atlas: AGLT2, MWT2_IU, BNL, OU_OSCER
    - OSG: Clemson, NERSC-Carver

Grid Applications

Derrick Stolee UNL Math Graph Theory Existence	7.4M cpu hrs	Condor File Transfer
Bob Powers UNL Chemistry Protein Active Site Structures	126880 jobs; 68k hrs	http (wget) LVS
Shi-Jian Ding UNMC OMSSA (Nuclear Matrix Proteomics)	270k hrs 2 wkends	SRM Squid 100 MB db
Jeff Thompson Digital Media Artistic Permutations	1500 hrs 1 afternoon	Condor File Transfer SRM

General Requirements for HCC

1. Worker Node OS = RHEL 5
   - CentOS5, SL5 OK!
   - We run/test on CentOS 5.
2. Worker nodes with internet access
   - Can be NAT.
   - Cannot just be http proxy.
3. OSG Certs
   - Required for GlideinWMS

General Requirements for HCC

4. OSG WN-Client

   – Should already be everywhere.

   – We use lcg-cp (with SRM) a lot.

5. Worker node local disk (flexible)

   – Assume 10GB per slot

   – In practice, we've never used more than 500MB

General Requirements for HCC

6. Worker node memory (flexible)

   – Assume 1.5-2GB per slot

   – 1 GB minimum

   – We detect memory with GlideinWMS

7. Squid Cache (Optional)

   – Or any http proxy

   – We use it for data 10MB < X < 150MB

Notes from Usage

- Squid usage at BNL:
  - Non-resolvable squid: squid.sec.bnl.local:3128

- We're fine with Preemption
  - Our target job length is 1 hour.

- We can use glExec if you want
  - Already run it at FNAL

Questions?